Cognitive Health Analysis in an Aging Population

1.0 Introduction

Cognitive health in aging populations is a growing area of concern, with various factors including demographics, SES, and biological measures playing potentially significant roles. This analysis aimed to explore these dimensions, focusing on their interaction and impact on cognitive health outcomes.

1.1 Research Questions and Hypotheses

We posed two critical research questions:

- How do demographic factors, socioeconomic status, and MRI measures interact to influence cognitive health outcomes, as measured by CDR and MMSE scores, in individuals?
- Specifically, is there an interaction effect between these variables on cognitive health outcomes?

Correspondingly, our hypotheses were framed as follows:

- H0 (Null Hypothesis): There is no interaction effect between demographic factors, SES, and MRI measures on cognitive health outcomes.
- HA (Alternative Hypothesis): There is a significant interaction effect between demographic factors, SES, and MRI measures on cognitive health outcomes.

2.0 Exploratory Data Analysis (EDA)

2.1 Data Cleaning and Preparation

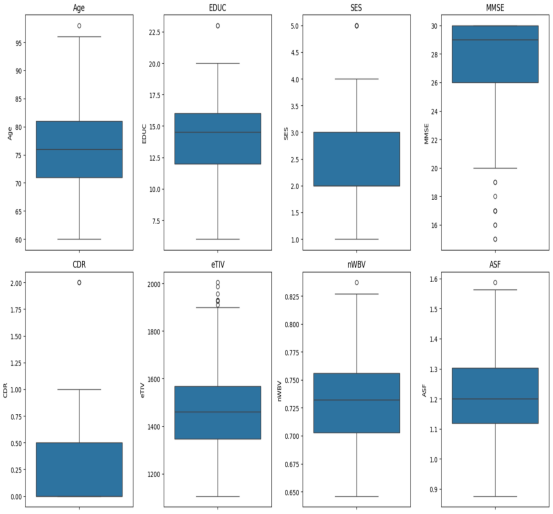
The dataset's initial examination revealed missing values in the SES and MMSE variables. The dataset contains a few missing values (Shown in figure 1): 15 in the SES variable and 1 in the MMSE variable. Given the relatively small proportion of missing data, we replaced the null value with the median value to maintain the dataset's integrity.

```
Subject ID     0
MRI ID         0
Group          0
Visit          0
MR Delay       0
M/F            0
Hand           0
Age            0
EDUC           0
SES           15
MMSE           1
CDR            0
eTIV           0
nWBV           0
ASF            0
dtype: int64
```

Figure 1: Count of the null values

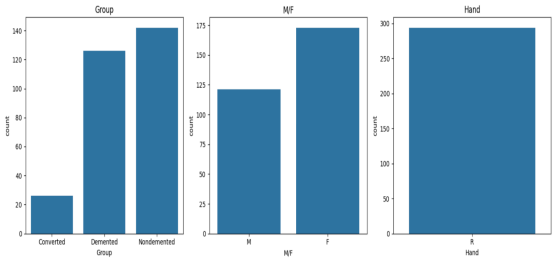
2.2 Exploratory Plots

Box plots for continuous variables shown in figure 2 in the next page, like Age, EDUC, MMSE, CDR, eTIV, nWBV, and ASF indicate the distribution and potential outliers within these variables. Most variables appear to have a relatively normal distribution, with some exceptions indicating potential outliers (e.g., EDUC, SES, MMSE).



The Group variable on the lower left of figure 3, which likely represents the cognitive health status (e.g., Demented, Nondemented), shows a balanced distribution among different categories. The M/F variable in the middle indicates a distribution of male and female participants, essential for considering gender as a factor in our analysis. The Hand variable shown on the right of figure 3 shows that the dataset predominantly includes right-handed individuals, which might reflect general population trends but could also influence the analysis if hand dominance affects the variables of interest.

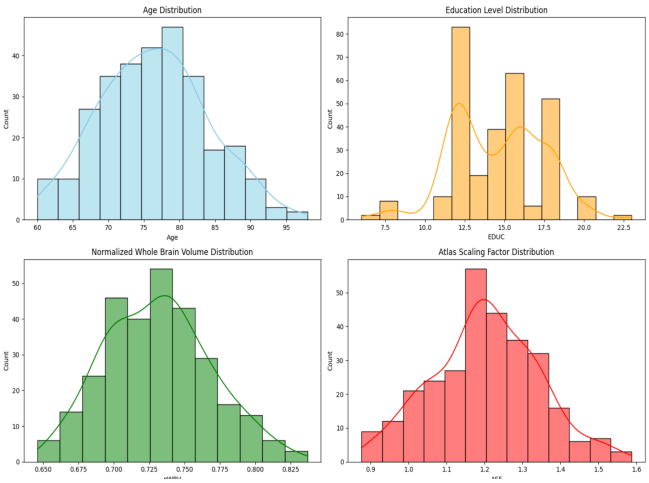
Figure 2: Box plots for continuous variables



The distribution plots for continuous variables provide insights into the data characteristics in figure 4:

- The Age Distribution plot shows a relatively normal distribution of ages among the participants, with slight skew towards older ages.

Figure 3: The distribution of categorical variables (Group, M/F, Hand)



- The Education Level Distribution reveals that most participants have around 12 to 16 years of education, suggesting a sample with a relatively high level of educational attainment.

- The Normalized Whole Brain Volume (nWBV) Distribution indicates a left-skewed distribution, where most participants have higher brain volumes, but there's a tail of individuals with lower volumes.

Figure 4: The distribution plots for continuous variables

- The Atlas Scaling Factor (ASF) Distribution appears to be right-skewed, indicating variability in the scaling factor applied to the brain images.

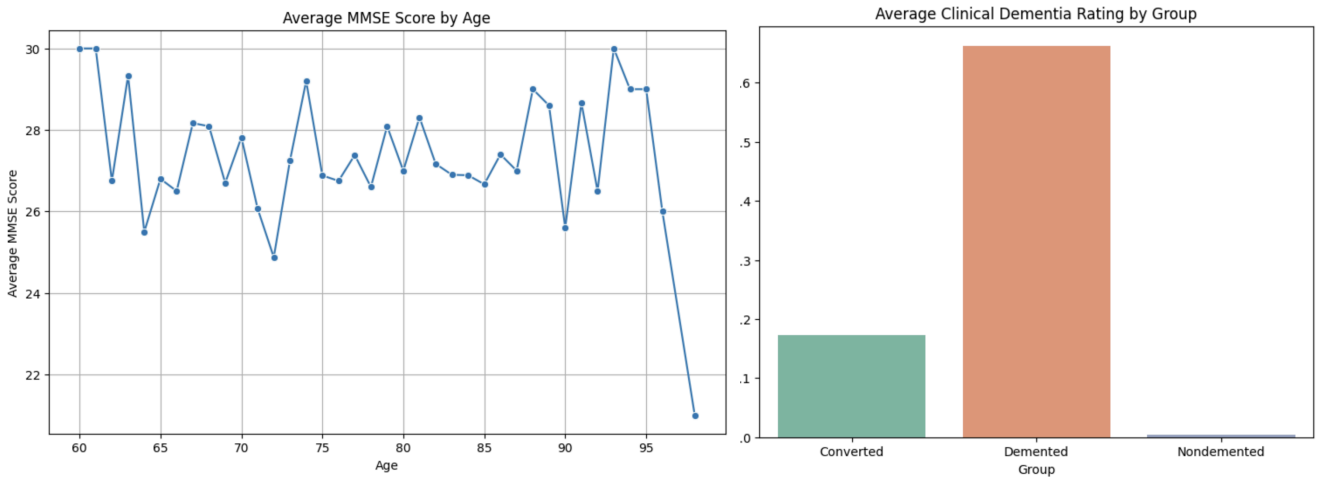


Figure 5: line chart of average MMSE scores (left) and bar chart of average Clinical Dementia Rating (right)

The line chart of average MMSE scores across different ages shows a trend where MMSE scores slightly fluctuate with age, potentially indicating variations in cognitive function as people age. The exact relationship and potential decline or variance would require further detailed analysis.

The bar chart of average Clinical Dementia Rating (CDR) by Group clearly illustrates differences in cognitive health status among the groups, with the Demented group having a higher average CDR, indicative of more significant cognitive impairment, compared to the Nondemented and Converted groups.

3.0 Mixed-Effects ANOVA Results

Our analysis conducted a mixed-effects model to explore the impact of various factors on CDR scores, which highlighted: A significant effect of dementia status on CDR scores, and the substantial influence of nWBV on cognitive health, with lower brain volumes associated with higher CDR scores.

Model:	MixedLM	Dependent Variable:	CDR
No. Observations:	294	Method:	REML
No. Groups:	150	Scale:	0.0290
Min. group size:	1	Log-Likelihood:	23.8412
Max. group size:	2	Converged:	Yes
Mean group size:	2.0		

	Coef.	Std.Err.	z	P>|z|	[0.025	0.975]
Intercept	0.886	0.559	1.586	0.113	-0.209	1.981
Group[T.Demented]	0.486	0.055	8.869	0.000	0.378	0.593
Group[T.Nondemented]	-0.156	0.053	-2.939	0.003	-0.261	-0.052
Age	-0.000	0.002	-0.169	0.865	-0.005	0.004
EDUC	0.007	0.007	0.912	0.362	-0.008	0.021
SES	0.000	0.019	0.013	0.990	-0.036	0.037
eTIV	0.000	0.000	0.271	0.787	-0.000	0.000
nWBV	-1.120	0.500	-2.242	0.025	-2.099	-0.141
Group Var	0.016	0.030				

The coefficient for the Demented group is significantly positive (coef = 0.486, $p < 0.001$) as shown in figure 6, indicating that being in the Demented group is associated with a higher CDR score compared to the reference group. Conversely, the Nondemented group has a significantly negative coefficient (coef = -0.156, $p = 0.003$), suggesting lower CDR scores compared to the reference group. These results highlight the substantial impact of dementia status on cognitive health as measured by CDR.

Figure 6: the mixed-effects ANOVA result

Among the continuous variables,the most demographic factors (Age, EDUC, SES, eTIV) did not show a significant association with CDR scores, suggesting that within the context of this model, they do not independently predict cognitive health outcomes. However, the normalized whole brain volume (nWBV) showed a significant negative association with CDR scores (coef = -1.120, p = 0.025), indicating that lower brain volumes are associated with higher CDR scores, which could be indicating that lower brain volumes are associated with higher CDR scores.

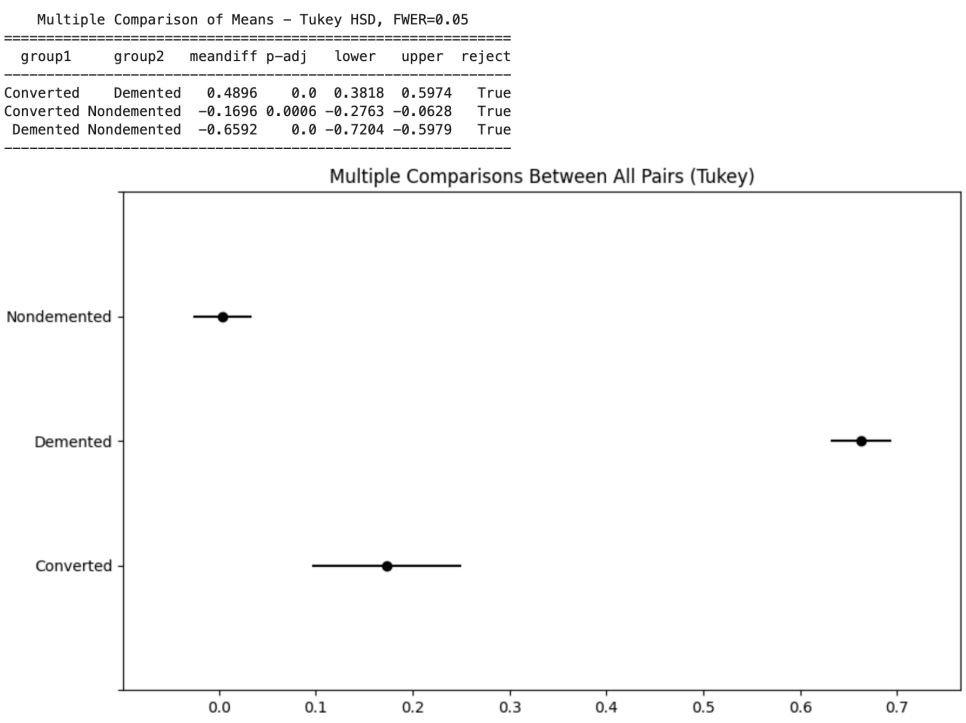


Figure 7: Post-Hoc Test

The Tukey HSD (Honestly Significant Difference) post-hoc analysis (in figure 7) reveals significant differences between all pairs of groups (Converted, Demented, Nondemented) in terms of their average Clinical Dementia Rating (CDR), with all comparisons showing a p-value of less than 0.05, indicating statistical significance:

- Converted vs. Demented: The average CDR for the Demented group is significantly higher than for the Converted group by 0.4896, suggesting more pronounced cognitive impairment in the Demented group.
- Converted vs. Nondemented: The average CDR for the Converted group is higher than for the Nondemented group by 0.1696, indicating a statistically significant difference in cognitive health status, with the Converted group showing more cognitive impairment.
- Demented vs. Nondemented: The Demented group has a significantly higher average CDR compared to the Nondemented group by 0.6592, reinforcing the clear distinction in cognitive impairment levels between these groups.

The pairwise comparison plot visually illustrates these differences, clearly showing the non-overlapping confidence intervals between the groups, which supports the rejection of the null hypothesis for each pair, indicating significant differences in cognitive health as measured by CDR.

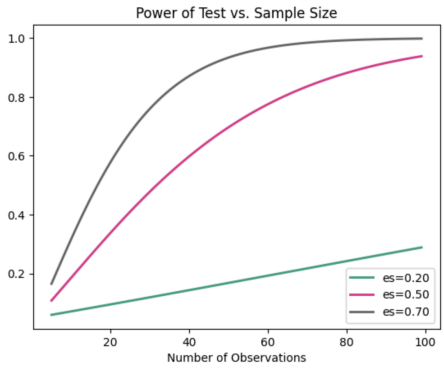


Figure 8: Power Analysis

The power analysis plot shown in figure 8 for t-tests shows the relationship between sample size and the power of the test across different effect sizes (0.2, 0.5, 0.7) with a significance level (alpha) of 0.05. For the specified parameters—a large effect size of 0.7, an alpha of 0.05, and a desired power of 0.91—the calculated appropriate sample size for each group in a theoretical experiment is approximately 46 participants.

This means that to achieve a power of 0.91 for detecting a significant effect (if one exists) with an effect size of 0.7 and at a significance level of 0.05, we would need at least 46 participants in each group of the study.

In conclusion, we reject the null hypothesis and there is a significant interaction effect between demographic factors, SES, and MRI measures on cognitive health outcomes. This alternative hypothesis is supported by the significant findings of the analysis, especially the interaction between the dementia status and nWBV with the cognitive health outcomes measured by CDR scores.

The analysis revealed a significant effect of dementia status on cognitive health outcomes. Participants identified as 'Demented' had higher CDR scores, indicative of more severe cognitive impairment, compared to 'Nondemented' individuals. This directly relates to our research question by confirming that the clinical grouping, which could be seen as a proxy for dementia progression, significantly interacts with cognitive health outcomes. The ANOVA results showed that lower normalized whole brain volumes (nWBV) were significantly associated with higher CDR scores. This finding provides evidence that MRI measures, specifically brain volume, are important biological factors related to cognitive health status.

While the model did not demonstrate a significant effect of age and education on CDR scores within this analysis, the overall model indicated that the group effects and nWBV are significant factors. This suggests that while demographic factors are important to consider, the biological measures and dementia status may have more direct associations with cognitive health outcomes in this dataset.The ANOVA analysis answered the research question by demonstrating dementia status and MRI measures like nWBV, significantly interacting with cognitive health outcomes. These findings suggest that interventions aimed at maintaining brain volume or mitigating the effects of dementia could be pivotal in preserving cognitive health in aging populations.